Misconceptions about Statistics in an Industrial Setting

William A. Brenneman
Statistics Department
The Procter and Gamble Company

Common Mistakes/Misconceptions in Statistics

1. If two group means are not statistically significantly different they must be equivalent
2. The R^2 is the *Holy Grail* in statistical modeling (Whole Kitchen Sink Mentality)
3. Confusion between confidence intervals, prediction intervals and tolerance intervals
4. Observing a strong correlation between two variables automatically implies cause and effect
5. Observing zero defects for a quality attribute in a sample implies there are zero defects in the population
6. Taking multiple measurements on an experimental unit (in a DOE setting) and treating them as if they are independent observations
7. Focus on parameter estimate of main effect when the factor actually interacts with another factor
8. Confusion about what type of protection a MIL-STD or ANSI sampling plan provides
9. One-at-a-time experimentation (miss interaction effects; create sub-optimal processes and formulations)
10. Learning a software package = Learning statistics

Misconception about Sampling Plans

Sampling plans for attributes (pass/fail) are often developed through the MIL-STD or ANZI/ASQC Z1.4 table look ups…or a nifty slide rule.

Example: Quality Assurance is developing a sampling plan for physical evaluation. An AQL (Acceptable Quality Level) of 1.0 is assigned to this particular type of defect. The goal is to show with high confidence the defect rate *is less than* the AQL. Lot size is 100,000.

Misconception #1 A MIL-STD or ANZI sampling plan provides this assurance.

Misconception about Sampling Plans

Table II-A—Single sampling plans for normal inspection (Master table)

(See 9.4 and

Sample size code letter	Sample size	0.010	0.015	0.025	0.040	0.065	0.10	0.15	0.25	0.40	0.65	1.0	1.5	2.5	4.0	6.5	10	15	25	40	65	100	150	250	400	650
		Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re	Ac Re
A	2	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↓	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22
B	3	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑↓	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31
C	5	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑↓	↑↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31	44 45
D	8	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑↓	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31	44 45	↑
E	13	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑↓	↑↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31	44 45	↑	↑
F	20	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑↓	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑	↑	↑	↑	↑
G	32	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑↓	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑	↑	↑	↑	↑	↑
H	50	↓	↓	↓	↓	↓	↓	↓	0 1	↑↓	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑	↑	↑	↑	↑	↑	↑
J	80	↓	↓	↓	↓	↓	↓	0 1	↑↓	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑	↑	↑	↑	↑	↑	↑	↑
K	125	↓	↓	↓	↓	↓	0 1	↑↓	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑	↑	↑	↑	↑	↑	↑	↑	↑
L	200	↓	↓	↓	↓	0 1	↑↓	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
M	315	↓	↓	↓	0 1	↑↓	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
N	500	↓	↓	0 1	↑↓	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
P	800	↓	0 1	↑↓	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
Q	1250	0 1	↑	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
R	2000	↑	↑	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑

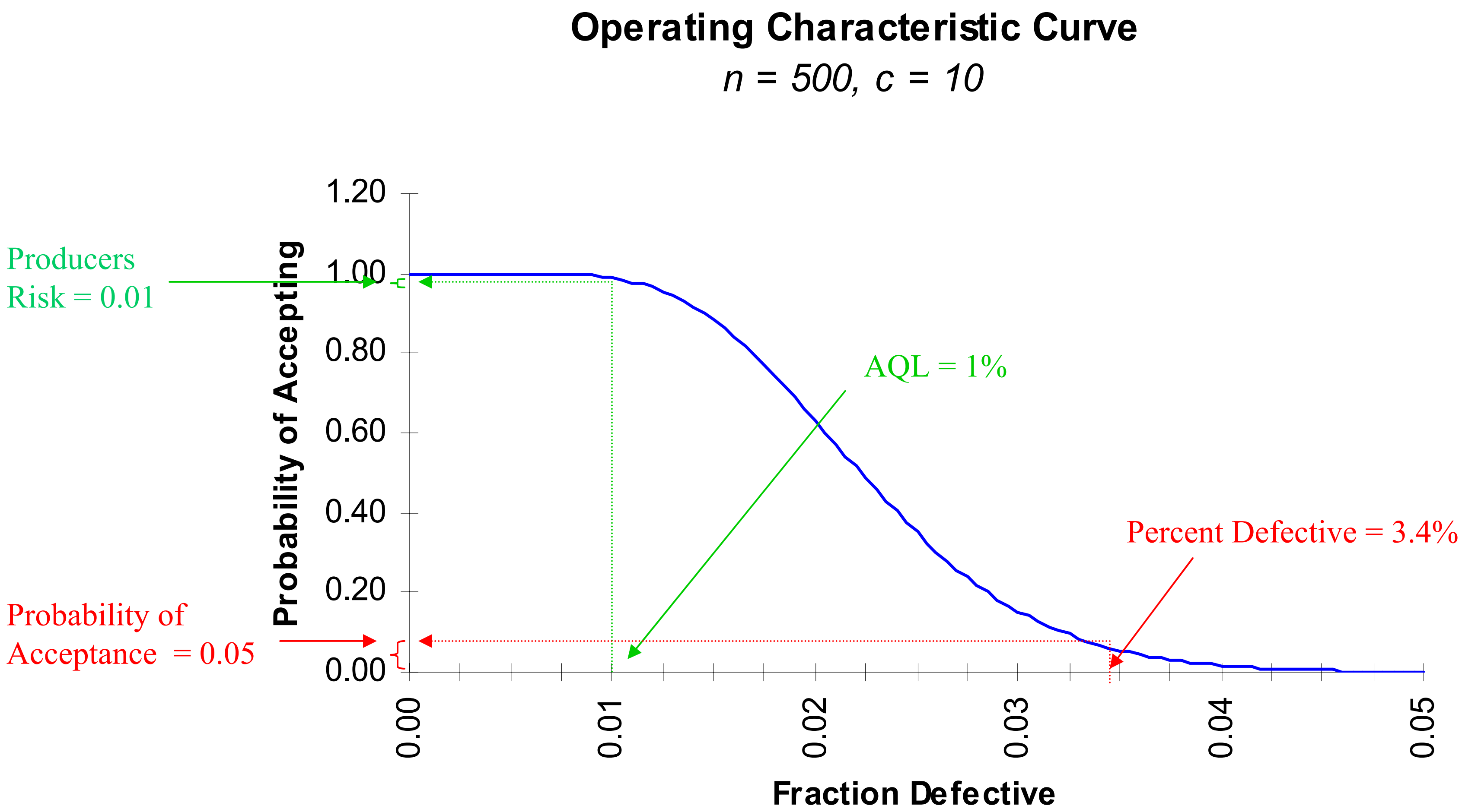
↓ = Use first sampling plan below arrow. If sample size equals, or exceeds, lot or batch size, do 100 percent inspection.
↑ = Use first sampling plan above arrow.
Ac = Acceptance number.
Re = Rejection number.

11

SINGLE NORMAL PLANS

In the ANSI/ASQC Z1.4, under normal inspection and an AQL = 1.0 , the sampling plan calls for a sample size of 500 and an acceptance number of 10.

Misconception about Sampling Plans



Plan provides high confidence that the defect rate is less than 3.4%, not 1%.

Misconception about Sampling Plans

Misconception #2 When I sample from a population, the resulting sample defect rate is equal to the population defect rate.

Example: 10 defects out of 500 samples means population defect rate = 10/500 = .02.

We handle these two misconceptions through a hands on sampling activity – similar to Deming's

Misconception about Sampling Plans

1. Sample 15 beads and record the number of RED beads found in the sample.
2. The acceptance number is 4. Record the decision of each sample.

	Black Bag			Red Bag	
Sample	# Red Beads	Decision		# Red Beads	Decision
1	4	Pass		6	Fail
2	2	Pass		9	Fail
3	2	Pass		7	Fail
4	3	Pass		6	Fail
5	3	Pass		6	Fail
6	6	Fail		7	Fail
7	2	Pass		5	Fail
8	2	Pass		4	Pass
9	2	Pass		6	Fail
10	2	Pass		10	Fail
11	2	Pass		8	Fail
12	0	Pass		3	Pass
13	1	Pass		2	Pass
14	5	Fail		7	Fail
15	3	Pass		7	Fail
16	3	Pass		7	Fail
17	3	Pass		9	Fail
18	1	Pass		6	Fail
19	1	Pass		9	Fail
20	4	Pass		5	Fail

Summary of Class Results

Black Bag % Red Beads = 17% 18%

Black Bag Reject % = 10% 10%

Red Bag % Red Beads = 43% 40%

Red Bag Accept % = 15% 20%

Truth

One at a Time Experiments

Reasons given to do 1 at a time…

1. The only way to know what is going on is to change one factor at a time (high school science class)
2. Much easier to understand
3. Takes less time
4. I know where we need to go anyway, why take the time to build a model

“Luckily”, very few of our scientists and engineers read Technometrics, August 2006, “Adaptive One-Factor-at-a-Time Experimentation and Expected Value of Improvement”

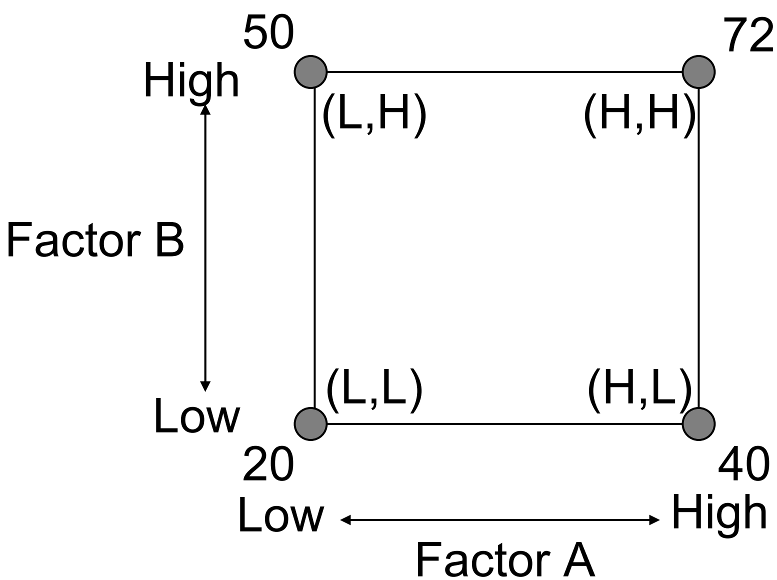
DOE Compared to One-at-a-Time Experiments

Advantages of DOE

1. More efficient (gets more information out of the same number of runs)
2. Allows estimation of interactions
3. Ensures that optimal settings will be found
4. Prevents confounding from occurring

Efficiency of DOE

Compared to One-at-a-Time Experiments



Factor A Main Effect

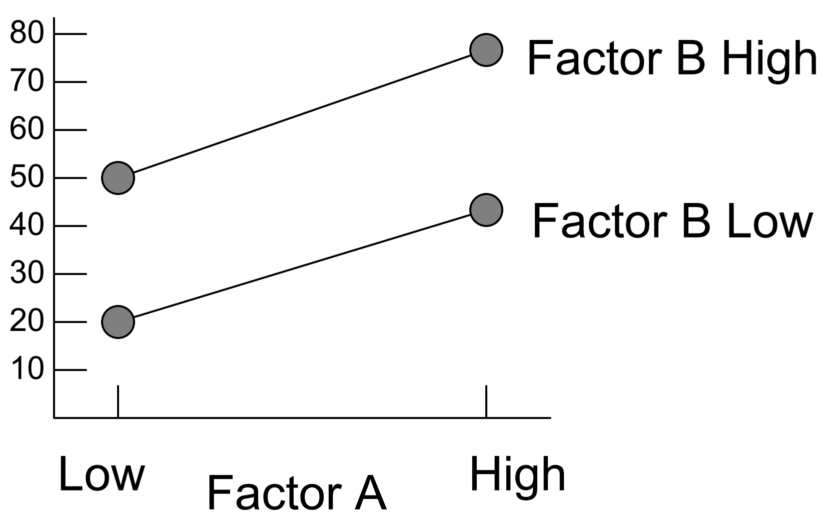
$$\frac{72+40}{2} - \frac{50+20}{2} = 21$$

Factor B Main Effect

$$\frac{72+50}{2} - \frac{40+20}{2} = 31$$

Factor AB Interaction Effect

$$\frac{72+20}{2} - \frac{50+40}{2} = 1$$



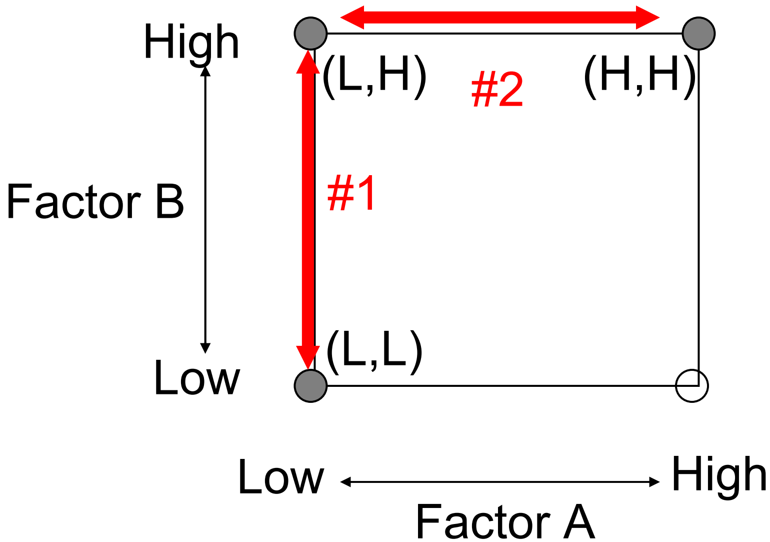
Effect Tests

Avg. of 2 data points vs Avg. of 2 data points

Total of 4 data points in design

Efficiency of DOE

Compared to One-at-a-Time Experiments

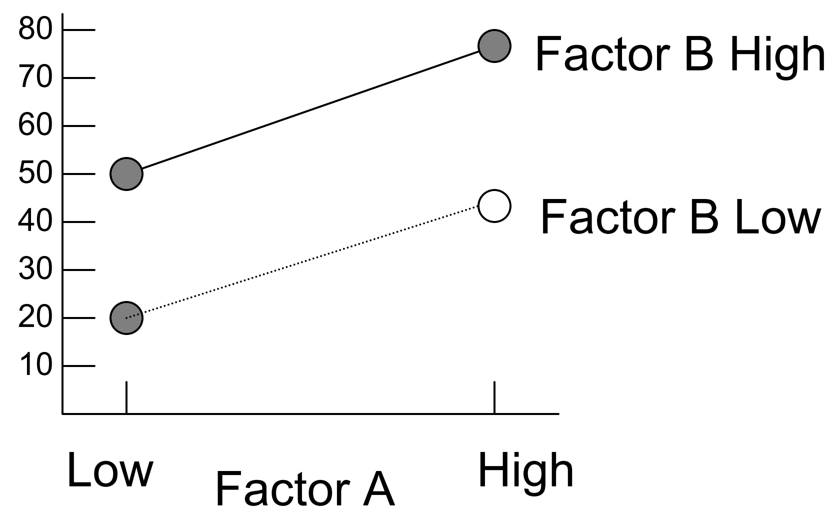


Experiment #1 – 2 data points (Low,Low) and 2 data points (Low,High)

- Compare (Low,Low) and (Low,High)
- Assume (Low,High) is better…

Experiment #2 – 2 data points (High,High)

- Compare (Low,High) and (High,High)



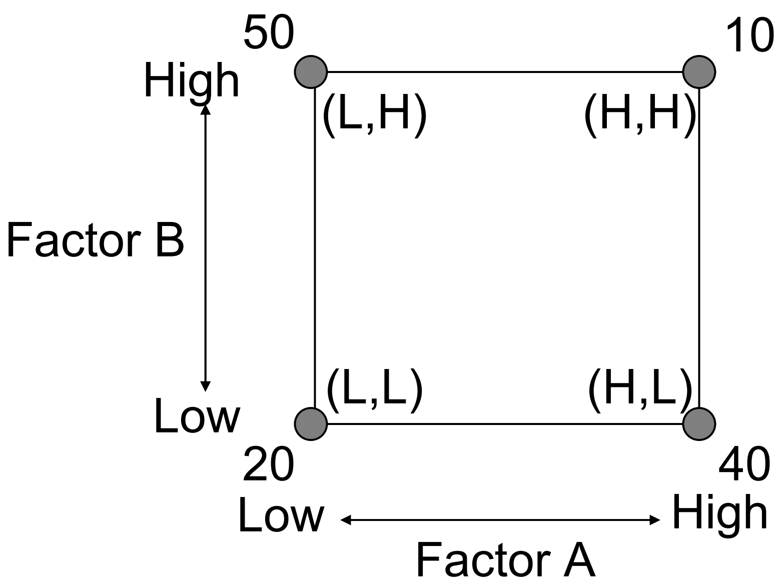
Main Effect Tests

Avg. of 2 data points vs Avg. of 2 data points

Total of 6 data points in this experiment

No Way to Estimate Interaction!

Ability of DOE to Estimate Interactions



Factor A Main Effect

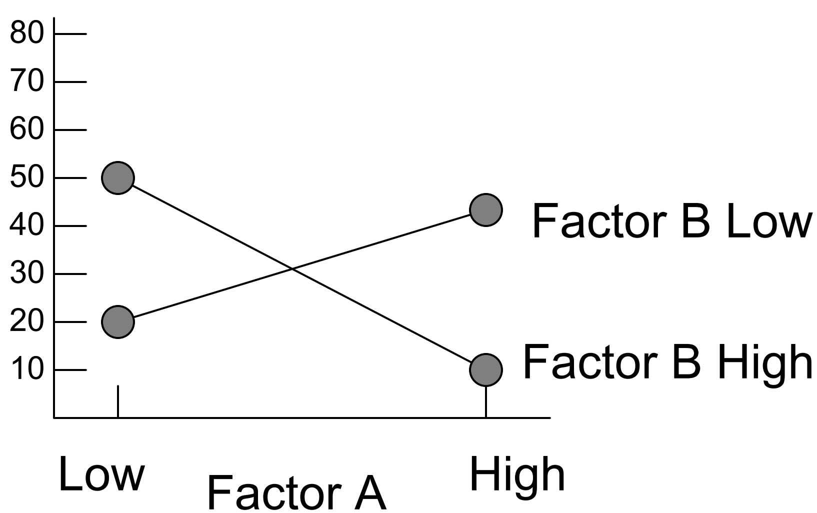
$$\frac{10+40}{2}-\frac{50+20}{2}=10$$

Factor B Main Effect

$$\frac{10+50}{2}-\frac{40+20}{2}=0$$

Factor AB Interaction Effect

$$\frac{10+20}{2}-\frac{50+40}{2}=30$$

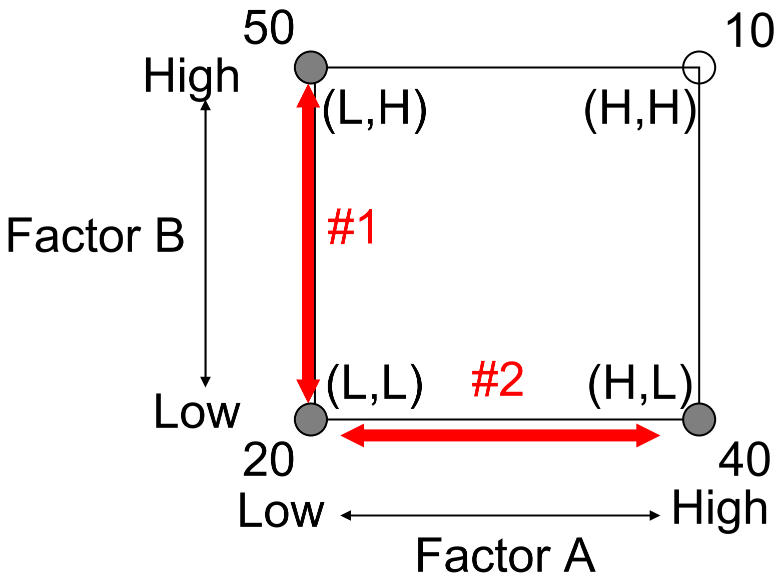


Strong Interaction Present

The effect of Factor A depends on the level of Factor B

The effect of Factor B depends on the level of Factor A

Optimization Problems with One-at-a-time Approach



Goal: Smaller is Better

Experiment #1

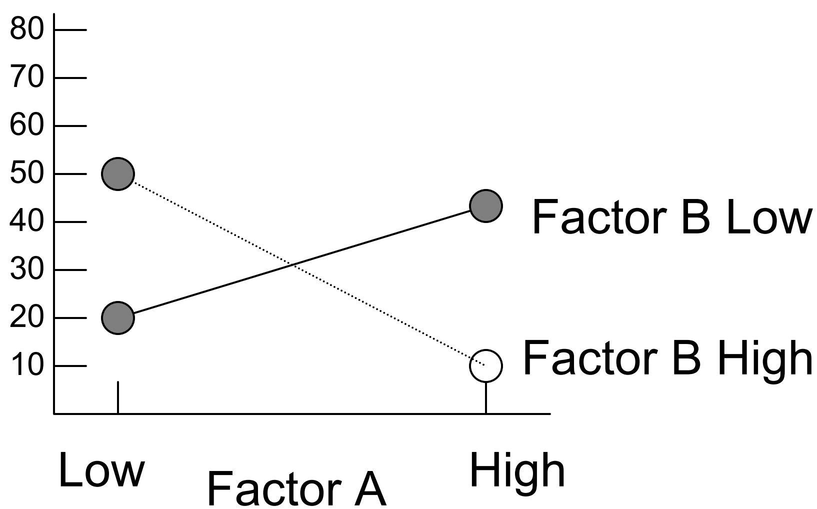
Compare (Low,Low) vs (Low,High)

→ (Low,Low) is better

Experiment #2

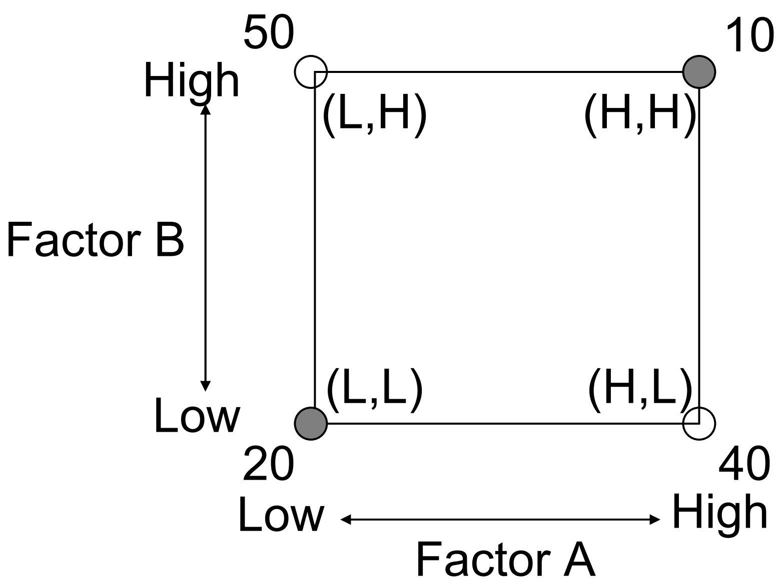
Compare (Low,Low) vs (High,Low)

→ (Low,Low) is BEST! NO!



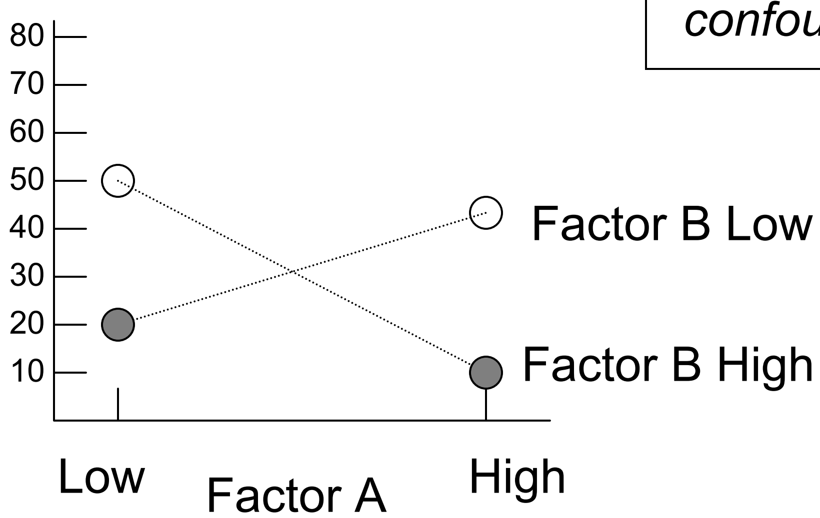
In this case the one-at-a-time-approach provides the ***wrong result*** because we missed an interaction effect

Advantage of DOE - Avoid Confounding Factors



In this experiment Factor A and Factor B both go from Low to High at the same time. No way to tell which factor caused the change.

These two factors are said to be *confounded* with one another.



Learning a Statistics Software Package ==

Learning Statistics

We have a corporate license for JMP, but same comments would occur with any other user friendly statistics package.

“Do you provide JMP training”
“I did a JMP analysis”
“I did a custom design”
“I was told that I should learn JMP”

Example – “I heard that JMP does response surface modeling so I tried it out on my data set!”

Learning a Statistics Software Package == Learning Statistics

Example – “I heard that JMP does response surface modeling so I tried it out on my data set!”

Design – 2^3 with center runs.

Parameter Estimates

Term		Estimate	Std Error	t Ratio	Prob>|t|
Intercept		-1.150824	0.557297	-2.07	0.0528
X1		0.3016151	0.197034	1.53	0.1423
X2		-0.139983	0.197034	-0.71	0.4861
X3		-0.129212	0.197034	-0.66	0.5198
X1*X2		-0.238069	0.197034	-1.21	0.2418
X1*X3		0.1414563	0.197034	0.72	0.4815
X2*X3		0.1618562	0.197034	0.82	0.4216
X1*X1	Biased	1.4841937	0.591102	2.51	0.0212*
X2*X2	Zeroed	0	0	.	.
X3*X3	Zeroed	0	0	.	.

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-1.150824	0.557297	-2.07	0.0528
X1	0.3016151	0.197034	1.53	0.1423
X2	-0.139983	0.197034	-0.71	0.4861
X3	-0.129212	0.197034	-0.66	0.5198
X1*X2	-0.238069	0.197034	-1.21	0.2418
X1*X3	0.1414563	0.197034	0.72	0.4815
X2*X3	0.1618562	0.197034	0.82	0.4216
X1*X1	1.4841937	0.591102	2.51	0.0212*

Learning a Statistics Software Package ==
Learning Statistics

"Statistical software will no more make one a statistician than would a scalpel turn one into a neurosurgeon."
-- Good & Hardin

Statistics Training at P&G – focus on a few key concepts and show the users how to use the software as a tool

Statistics Training + Collaboration = Increased Capabality

Questions

Contact Information

brenneman.wa@pg.com
513-622-3195